모멘트 특성을 이용한 다중 객체 이미지 검색 시스템 구현

안광일[†] · 안재형[††]

요 약

영상과 같은 다양하고 복잡한 데이터 검색은 기존의 키워드를 이용한 검색이 아닌 내용 기반 검색 방법이 요구된다. 본 논문에서는 물체의 위치이동이나 회전, 크기변화 등과 같은 각종 변환에 민감하지 않은 불변모멘트(invariant moments)값의 특성을 이용하여 사용자 질의로서 입력된 객체를 효율적으로 검색할 수 있는 시스템을 구현하였다. 영상내의 단일 객체뿐만 아니라 다중 객체들도 효과적으로 검출하기 위해 레이블링(labeling) 알고리즘을 적용해 각각의 객체를 따로 분리하여 불변모멘트를 적용하는 방법을 이용했다. 또한, 검색 시간 단축 및 영상의 효율적인 인덱싱(indexing)을 위해 해싱을 응용한 기법을 적용하였다. 실험결과, precision 85%, recall 23%의 높은 검색효율을 보였고 기존의 전체 영상의 특징을 가지고는 정확히 표현할 수 없는 객체들의 모양을 정확히 표현해줌으로써 보다 정확한 검색 결과를 얻을 수 있었다.

Implementation of System Retrieving Multi-Object Image Using Property of Moments

Kwang-Il Ahn[†] and Jae-Hyeong Ahn[††]

ABSTRACT

To retrieve complex data such as images, the content-based retrieval method rather than keyword-based method is required. In this paper, we implemented a content-based image retrieval system which retrieves object of user query effectively using invariant moments which have invariant properties about linear transformation like position transition, rotation and scaling. To extract the shape feature of objects in an image, we propose a labeling algorithm that extracts objects from an image and apply invariant moments to each object. Hashing method is also applied to reduce a retrieval time and index images effectively. The experimental results demonstrate the high retrieval efficiency i.e precision 85%, recall 23%. Consequently, our retrieval system shows better performance than the conventional system that cannot express the shape of objects exactly.

1. 서 론

현재 우리가 접하는 멀티미디어정보 중 영상과 같이 다양한 형태의 데이터들은 그것과 연관된 텍스트나 키워드를 사용하는 기존의 검색방법으로는 그 영상을 제대로 묘사해 주지 못한다. 이들 검색 방법은 각 영상에 입력된 초기 텍스트 정보에 전적으로 의존하기 때문에 만약 질의어가 초기에 묘사되지 않은 영상속성을 참조하게 되면 검색은 대부분 실패하게 된다. 따라서 칼라(color)나 모양(shape), 질감(texture) 등의 성분을 이용한 내용기반 검색방법이 필요하게 되었다. 그 동안 이러한 성분들을 이용한 내용기반 검색방법에 대해 다양한 연구가 있어 왔다. Hirata는 사용자에 의한 스케치 입력이나 칼라나 복사본 영상으로 데이터를 검색할 수 있는 방법을 제시하였다[1]. IBM사의 QBIC 시스템에서는 입력으로 텍스트, 스

† 준회원, 충북대학교 정보통신공학과 석사과정
†† 정회원, 충북대학교 정보통신공학과 교수

케치, 칼라정보, 레이아웃(layout)이나 구조적 묘사 등의 정보를 기반으로 한 대규모 데이터베이스 질의가 가능한 검색방법을 개발하였다[2]. 버클리 소재 캘리포니아 대학에서 수행한 chabot 프로젝트에서는 애트리뷰트와 특징을 복합적으로 사용한 검색기법을 제안하였고, Virage 시스템에서는 검색에 사용되는 특징들을 개발자가 응용에 맞게 추가하거나 삭제할 수 있는 기능을 제공하였다[3]. 이러한 기존의 내용기반 검색시스템에 대한 연구들은 영상내의 칼라 히스토그램 등을 이용한 칼라정보와 곡률, 모멘트 등의 모양특징 정보, 그리고 질감 등을 이용한 검색 방법이 대부분을 차지하고 있으나, 이들 정보들만을 가지고 광범위한 영상들을 정확하게 표현해줄 수 없다. 특히 모양특징 검색에서는 사용자가 검색하고자 하는 객체의 모양을 정확히 표현해준다는 것은 어려운 일이다. 이에 본 논문에서는 기존에 주로 다루어진 단일객체 검색시스템에서 확장하여 한 영상내에서 여러 개의 객체들도 독립적으로 표현해 줄 수 있는 레이블링 기법을 제안하였고, 각각의 객체들을 위치나 크기변화 등 각종 변환에 불변적인 특징을 갖는 불변모멘트를 이용하여 표현하였다. 이로써 사용자는 선택적으로 원하는 객체만을 검색할 수도 있고, 그 객체를 포함하는 모든 영상들을 검색할 수도 있는 기능을 갖게 되었다. 또한 DB 저장 및 검색방법에서 해싱기법을 기반으로 한 개선된 모델 재구성 방법을 제안하여 효율적인 검색시스템을 구현하였다.

2. 객체의 특징 추출

영상내 객체의 특징추출 절차는 그림 1과 같다. 전처리단계로서의 잡음제거 및 영상 이치화, 그리고 객체 레이블링 작업이 있고, 다음으로 불변모멘트를 이용한 특징값 추출이 있다. 전처리 단계에서의 이치화 작업은 객체들과 배경을 분리해주기 위함이다. 또한 초기 입력 영상은 조명등에 민감하기 때문에 객체를 정확히 분리하기 위해 히스토그램 균일화나 메디안 필터링 등과 같은 잡음 제거작업이 선행되어야 한다. 이치화 단계는 전체영상을 크게 두 그룹으로 나누고 그레이 레벨값에 따른 그룹간의 평균과 분산값을 이용하여 최적의 임계값을 얻는 Otsu의 방법을 이용하였다[4].

객체 레이블링 단계는 단일 객체가 아닌 여러 개의 객체가 하나의 영상에 존재할 경우 각각의 객체를 레이블화 하게 된다. 즉 세부 단계로서 이치화된 영상을 2차 미분연산자인 라플라시안 연산자를 이용하여 윤곽선을 검출하게 되고, 그림 2와 같은 8방향 체인코드(chain code)를 기반으로 한 3×3 mask법을 이용하여 윤곽선을 추적하게 된다[5]. 그 다음 검출된 각각의 객체들의 윤곽선에 대해 고유한 레이블값을 삽입하게 되며 레이블값에 따른 객체들의 전체분포값을 복원하게 된다. 이렇게 구분된 각각의 객체들은 불변모멘트를 적용하여 그 객체들의 특징값들을 추출하게 된다. 불변모멘트의 특징은 물체의 위치이동이나 회전, 크기 변화 등과 같은 각종 변환에 불변적인 성질이 있다는 것이다.
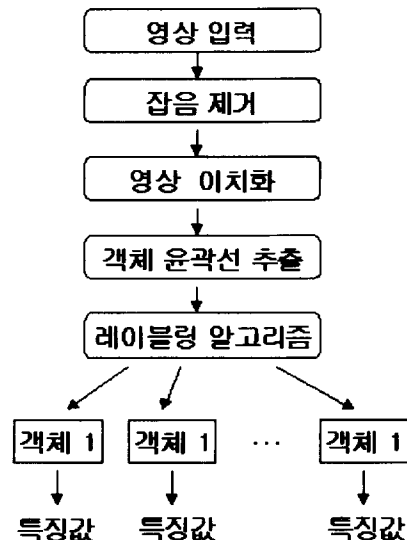
그림 1. 객체의 특징추출 과정
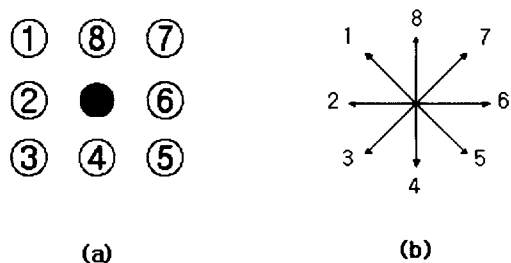
그림 2. 윤곽선 추적을 위한 8방향 마스크
(a) 3×3마스크 (b) 8방향벡터코드

이 논문에서는 이와 같은 객체의 특징값을 추출하기 위하여 객체에 대한 불변모멘트를 적용하였다[4,6,7]. 모멘트는 이미지 분석측면에서 보면 이미지 세그먼트(segment)의 특성을 나타내거나 분석적인 특징들을 추출하는데 사용된다. 이산값(discrete value)을 갖는 2차원 이미지 공간(i, j)에서 M×N 크기를 가진 영상에서의 각 픽셀 값을 f(i, j)라 하면, (p + q) 차 모멘트는 아래 식(1)과 같다.

$$m_{pq} = \sum_{i=0}^{M} \sum_{j=0}^{N} i^p j^q f(i,j) \tag{1}$$

만일 오직 한 객체만 존재하고, 배경값이 모두 '0' 이라면, 그 객체의 중심은

$$x_c = m_{10} / m_{00} \quad , \quad y_c = m_{01} / m_{00} \tag{2}$$

이 되는데, 여기서 0차 모멘트(m_{00})는 객체의 면적을 나타내고 1차 모멘트(m_{10}, m_{01})는 i, j축의 분포값이 된다. 객체의 중심이 구해졌으면 객체의 좌표값에서 중심을 뺀 값을 모두 p, q차수의 가중치에 따라 합산하여 중심모멘트 및 정규화된 중심모멘트를 다음과 같이 구할 수 있다.

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q f(x, y) \tag{3}$$

$$\eta_{pq} = \mu_{pq} / \mu_{00}^{(p+q+2)/2} \tag{4}$$

이 정규화된 중심모멘트를 이용해 각 객체에 대한 불변 모멘트는 다음과 같이 정의된다[7,8,9].

$$M_1 = \eta_{20} + \eta_{02}$$
$$M_2 = (\eta_{20} - \eta_{02})^2 + 4(\eta_{11})^2$$
$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2$$
$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \tag{5}$$

그림 3은 불변모멘트값의 특징을 설명하기 위한 예제이다. 영상내에서 하나의 같은 객체가 이동, 축소, 반사 등과 같은 각종 변환이 되었을 때, 이 객체에 대한 4개의 불변모멘트값을 보여주고 있는데 값들의 변화가 없이 거의 일정한 값을 가짐을 알 수 있다.

불변 모멘트값은 원래 7개의 값을 갖게 되는데 그 값들 중 M_5~M_7값은 높은 차수들의 모멘트 합으로 나타내어지므로 객체를 더 정밀하게 표현해 줄 수는 있지만 잡음에 너무 민감한 성질을 나타내게 된다[10]. 그래서 본 논문에서는 M_1~M_4의 4개 값을 선택적으로 검색에 이용하여 입력 영상 객체의 잡음에 의한 오류 검색을 감소시키게 되고, 계산량 감소에 의한 데이터베이스 인덱싱 시간을 줄일 수 있는 효과를 얻을 수 있었다.
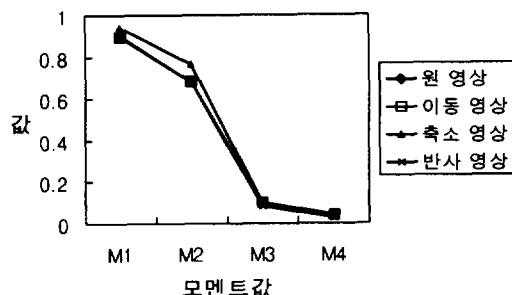
그림 3. 불변모멘트 값들의 특성

3. 해싱기법을 이용한 개선된 모델 재구성

해싱이란 찾고자 하는 데이터의 키(key)를 이용하여 데이터베이스에서 데이터가 저장되어 있는 위치를 직접 산술적으로 계산하여 원하는 데이터를 취하는 방식의 검색 방법이다. 본 논문에서는 모든 객체에서 얻어진 모멘트 값들의 거리를 계산하여 이들 중 유사성이 가장 높은 객체의 모멘트 값을 참조 객체값으로 선정하고 각각 객체의 모멘트 값으로부터 참조 객체값까지의 거리를 키값으로 하여 해시테이블을 구성하였다. 이때 키값의 범위는 참조 객체와의 거리에서 산출된 평균과 분산값을 기준으로 구성된다. 이 키값의 범위는 DB 영상을 각각의 클래스로 분류하는 기준이 된다. 참조 객체의 선정기준은 두 개의 객체 S_i, S_j의 모든 쌍(S)에 대한 거리를 유클리디언 거리 측정법을 이용하여 구한다. 이때 객체들의 총 수가 n 이면 상호대칭의 관계를 가지므로 n(n-1)/2 의 계산량을 필요로 한다. 그 다음으로 전 단계에서 구한 거리를 이용해 아래 식(6)의 조건을 만족하는 객체를 참조 객체(s_r)로 선정하게 된다.

$$s \in S, \ \max\{ Dist(s, s_r) | s_r \in S\} \text{ is Minimum} \tag{6}$$

질의 영상이 입력되었을 때 검색 방법으로는 질의 영상 객체와 참조 객체와의 거리 측정에 의해 구해진 거리값이 DB내의 어느 영상 클래스에 해당되는지를 판별하여 그 클래스 내에서 유사도 검색을 실시하

게 된다.

그림 4는 위에서 설명한 전체적인 해시테이블 작성과정을 보여주고 있는데, 이러한 방법으로 본 논문에서는 영상 데이터들이 참조객체와의 거리 계산에 의해서 구한 평균과 분산값에 의해 5개의 클래스로 나뉘어져 검색공간이 전체공간의 20%로 줄어드는 효과를 얻을 수 있었다.
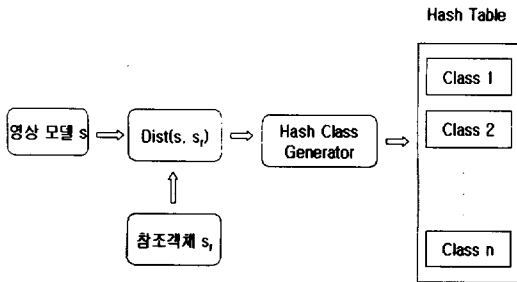
그림 4. 해시 테이블 작성 과정

4. 시스템의 구성

전체적인 시스템 구성은 그림 5와 같다. 원 영상들은 히스토그램 균일화 및 이치화 과정을 행한다. 여기에서 처리된 영상은 단일 객체인지 다중 객체인지를 판별하기 위한 단계로서 객체의 윤곽선을 검출하게 되고, 이것을 레이블링 알고리즘을 적용해 각 객체를 구별해 준다. 이렇게 구해진 객체들은 특징값을 추출하게 되는데, 이 값들은 제안된 해시 테이블 형태로 저장된다. 사용자 질의 영상 검색도 같은 과정으로 객체를 추출하게 되고 불변 모멘트를 적용하여 특징값들을 얻게 된다. 이 값들은 이미 DB에 해시테이블 형태로 저장되어 있는 객체 특징값들과 비교를 하게 되고 가장 유사성이 큰 이미지 순서로 화면에 보여주게 된다.
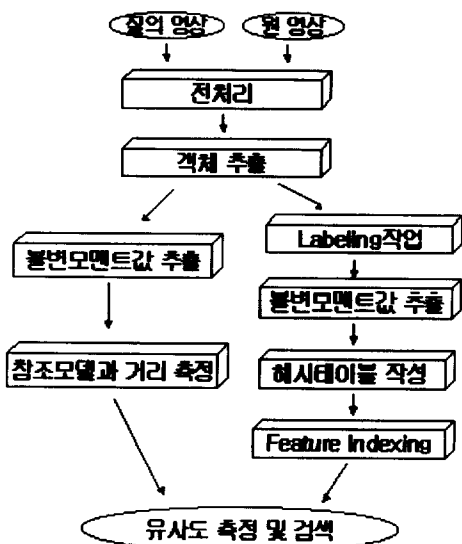
그림 5. 시스템 구성도

5. 실험 및 결과

본 실험은 Visual C++를 사용하여 프로그래밍 하였으며, 실험에는 CCD 카메라를 이용하여 얻은 JPEG 칼라영상을 340*240 해상도의 256 그레이레벨 비트맵(bitmap)영상으로 변환하여 사용하였다. 실험에서 모델로 사용되어진 물체는 드라이버, 망치, 니퍼, 롱로즈 등과 같은 공구 및 부품에 관련된 영상들이며, 이들을 각각 전처리 한 다음 모델링하고, 사용자 입력영상과의 유사도 측정에 의해 모델과 가장 유사한 영상을 찾아준다.

실험조건은 단일객체 영상 200개와 다중객체를 포함한 영상 300개를 대상으로 하였다. 질의 입력 방법은 DB내 영상 중에서 임의의 질의 영상 10개를 무작위로 선택하여 각각의 질의영상 객체들의 모멘트값을 구하고, 이를 DB내의 영상들과 유사도 검색을 실시하여 유사도가 높은 순으로 최대 6개까지 화면에 출력하게 된다. 내용기반 검색시스템은 정확한 매칭이 아닌 유사매칭을 수행하므로 시스템 평가 방법으로 normalized precision과 recall statistics를 근거로 하였다. 주어진 질의와 관련된 항목 수를 T라 하고, T로부터 검색된 항목의 총 수를 Tr, 검색된 항목 중 질의와 유사한 항목의 수를 Rr 이라고 할 때 precision은 Rr / Tr 이고 recall statistics는 Rr / T 이 된다.

표 1은 하나의 객체만 존재하는 영상들만을 대상으로 했을 때의 실험결과로서, 본 논문에서 구현된 시스템과의 성능 비교를 위해 IBM의 QBIC 시스템에 제안된 시스템과 같은 수의 질의 갯수 및 DB크기를 주고 검색을 실시했을 때 제안된 시스템의 검색결과와 비교한 것이다. QBIC 시스템에서의 질의를 위해 사용한 특징으로는 칼라 히스토그램과 질감 특징, 그리고 레이아웃 등이 있다. 이들 각각의 특징으

로 검색한 결과를 살펴보면 제안된 시스템은 precision 85 %, recall 42 %가 나온 반면, QBIC 시스템에서는 평균 40 %정도의 precision과 평균 20 %의 recall을 나타내었다.

표 2는 단일 객체를 갖는 영상들뿐만 아니라 다중 객체를 갖는 영상들도 포함하여 실험한 결과를 나타내고 있다. 여기서도 표 1에서와 같은 조건을 주고 QBIC 시스템과 제안된 시스템간의 검색 결과를 보여주고 있다. 결과값을 비교해 보면 QBIC 시스템에서의 precision과 recall은 평균 50 %와 15 %로 나온 반면, 제안된 시스템의 precision과 recall은 85 %와 23 %로 높은 검색효율을 나타내고 있다. 이러한 결과는 본 논문에서 제안하고 적용한 레이블링 알고리즘을 통해 각 객체들이 정확하게 표현되었기 때문에

표 1. 단일객체영상들의 질의에 따른 실험결과

system / param	QBIC System			Proposed System
Query Property	Color	Texture	Layout	Invariant Moments
DB size	200	200	200	200
Image Query수	10	10	10	10
T	120	120	120	120
Tr	60	60	60	60
Rr	26	24	23	51
Precision	43%	40%	38%	85%
Recall	22%	20%	19%	42%

표 2. 다중객체를 포함한 영상들의 질의에 따른 실험결과

system / param	QBIC System			Proposed System
Query Property	Color	Texture	Layout	Invariant Moments
DB size	300	300	300	300
Image Query수	10	10	10	10
T	224	224	224	224
Tr	60	60	60	60
Rr	30	33	32	51
Precision	50%	55%	53%	85%
Recall	13%	15%	14%	23%

가능했다. 또한 불변모멘트의 특성을 이용하여 다양한 형태의 객체를 효과적으로 표현하였기 때문에 QBIC 시스템보다 높은 정확도를 얻을 수 있었다.

그림 6에서는 사용자 입력 영상으로 검색을 하기 위한 어플리케이션을 나타내고 있다. 여기서 사용자는 모드 선택에 의해 찾고자하는 객체만를 가지고 있는 단일객체 영상들을 검색할 수도 있고, 그 객체를 포함하는 다중객체 영상들을 선택적으로 검색할 수도 있다. 이렇듯 사용자에 의한 선택적인 검색방법을 통해 검색의 시간 단축 및 효율성을 증대시키는 효과를 가져왔다.
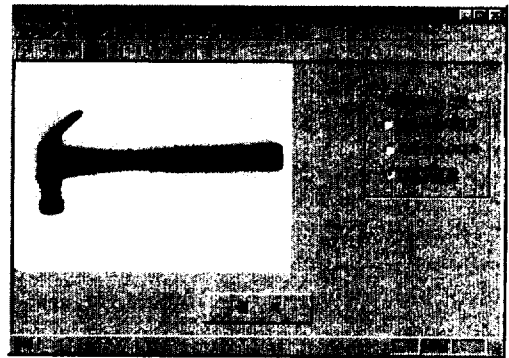
그림 6. 검색을 위한 사용자 인터페이스

그림 7은 질의 영상(망치)을 가지고 검색한 결과를 보여주고 있다. 질의 영상을 먼저 이치화하고 불변모멘트에 의한 특징값들을 구한 후, DB에 인덱싱되어있는 영상들의 특징값들과 비교하여 검색결과를 보여주게 된다. 검색결과는 좌측상단부터 가로로 유사도가 큰 순서대로 6개의 영상을 보여주고 있는데, 질의 영상의 객체를 포함하는 단일객체 영상 및 다중객체 영상들이 정확하게 검색되었음을 알 수 있다.

그림 7. 질의영상(망치) 검색결과

표 3은 그림 7의 검색결과 영상들의 모멘트값을 보여주고 있다. 결과 영상들 모두 약간의 오차를 제외하고는 거의 일정한 값의 분포를 가짐을 알 수 있다. 그림 8은 표 3의 결과 영상의 객체 모멘트 값들간의 오차값을 비교한 그래프이다.

표 3. 결과 영상의 불변모멘트값

특징값 / 결과 영상순위	M_1	M_2	M_3	M_4
1	0.896876	0.682405	0.101779	0.039106
2	0.915618	0.717464	0.110409	0.044576
3	0.896862	0.682619	0.083899	0.029290
4	0.898995	0.689497	0.102688	0.042634
5	0.902302	0.694691	0.100693	0.038843
6	0.927956	0.741985	0.093596	0.036185
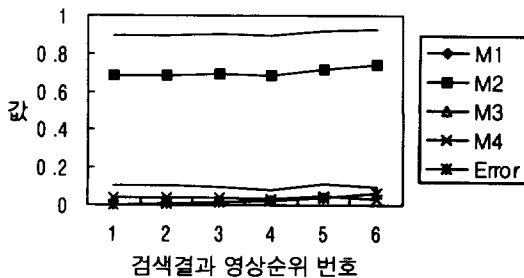
그림 8. 결과 영상의 성능 비교

6. 결 론

본 논문에서는 불변모멘트를 적용하여 각 영상의 객체들의 특징값을 구했고, 이 값들을 가지고 사용자가 원하는 객체를 효과적으로 검색하기 위한 시스템을 구현하였다. 물체의 회전, 이동, 크기 변화나 각종 변환에도 값의 변화가 거의 없는 특징을 갖고 있는 불변모멘트를 이용했으며, 기존의 검색 시스템을 보완하여 한 영상내에 여러 개의 객체가 존재하더라도 각각의 객체들을 정확하게 표현할 수 있는 레이블링 알고리즘을 적용하였다. 또한 검색의 시간 단축 및 효율을 개선하기 위해 해싱기법을 응용한 영상모델 재구성기법을 적용하여 보다 효율적인 시스템을 구현하였다.

이번 실험에서는 객체와 배경을 구분하는 영상의 이치화 과정에서 조명에 대한 처리가 미비한 관계로 몇 개의 영상에서 약간의 오차가 발생했다. 그러나 이 부분은 좀더 보완된 전처리 과정을 거친다면 해결될 수 있을 것으로 판단된다. 이번에 구현된 시스템은 공구 영상과 같이 배경과 객체가 뚜렷이 구분된 영상을 이용하였다. 앞으로 이 시스템을 보완하여 칼라 정보를 추가한다거나, 좀더 복잡한 형태의 객체도 분류할 수 있는 알고리즘을 적용한다면 여러 분야의 검색 시스템에서 활용이 가능하리라고 기대된다.

참 고 문 헌

[1] 鄭元一 외 3명, "내용기반 화상 검색시스템의 설계 및 구현," 電子工學會論文 第33券 B編 第7號, pp. 1080-1089, 1996
[2] R.Barber, W.Equitz, et al., "Query by content for large on-line image collections," A Guided Tour of Multimedia Systems and Applications, IEEE Computer Society press, pp. 357-378, 1995
[3] 박치항외 2명, "A Study on DBMS Technology for the 21 Century," 한국전자통신연구원, Vol. 2, pp. 575-615, 1997
[4] 이지용, 이병곤, "로버트와 화상처리를 이용한 위험기계의 loading 작업," 산업안전학회지, Vol. 9, No. 1, pp. 31-39, 1994
[5] Xiaolong Dai and Siamak Khorram, "A Feature-Based Image Registration Algorithm Using Improved Chain-Code Representation Combined with Invariant Moments," IEEE TRANSACTIONS, Vol. 37, No. 5, pp. 2351-2362, 1999
[6] Alexander G. Mamistvalov, "n-Dimensional Moment Invariants and Conceptual Mathe--matical Theory of Recognition n-Dimensional Solids", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, Vol. 20, No. 8, 1998
[7] Gouda I. Salama and A.Lynn Abott, "Moment Invariants and Quantization Effects," IEEE PROCEEDINGS, pp. 157-163, 1998
[8] Th.M.Hupkens and J.de Clippeleir, "Noise and

intensity invariant moments," Pattern Recog--nition Letters 16, pp. 371-376, 1995
[9] Lixin Shen and Yunlong Sheng, "Noncentral image moments for invariant pattern recognition," OPTICAL ENGINEERING, Vol. 34, No. 11, pp. 3181-3186, 1995
[10] David Slater and Glenn Healey, "Modeling the sensitivity of moment invariants in a recognition system," Optical Society of America, Vol. 15, No. 5, pp. 1068-1075, 1998

안 광 일

1999년 충북대학교 정보통신공학과 졸업(공학사)
1999년~현재 충북대학교 정보통신공학과 석사과정
관심분야 : 내용기반검색, 패턴인식, 영상처리
E-mail : aki720@hanmail.net

안 재 형

1981년 충북대학교 전기공학과 졸업(공학사)
1983년 한국과학기술원 전기 및 전자공학과(공학석사)
1992년 한국과학기술원 전기 및 전자공학과(공학박사)
1987년~현재 충북대학교 전기전자공학부 교수
관심분야 : 영상통신 및 영상정보처리, 멀티미디어 제작 및 정보제공, 인터넷 통신 및 프로그래밍
E-mail : jhahn@cbucc.chungbuk.ac.kr